

H4
PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,760

DATE: 08/30/2001

TIME: 09:01:49

Input Set : A:\00019900.txt

Output Set: N:\CRF3\08302001\I868760.raw

3 <110> APPLICANT: The Horticulture and Food Research Institute of New Zealand Limited
 5 <120> TITLE OF INVENTION: Serine Protease Inhibitor
 7 <130> FILE REFERENCE: 514274-2001
 9 <140> CURRENT APPLICATION NUMBER: 09/868,760
 C--> 10 <141> CURRENT FILING DATE: 2001-06-21
 12 <150> PRIOR APPLICATION NUMBER: PCT/NZ99/00227
 13 <151> PRIOR FILING DATE: 1999-12-23
 15 <150> PRIOR APPLICATION NUMBER: NZ 336906
 16 <151> PRIOR FILING DATE: 1999-07-23
 18 <150> PRIOR APPLICATION NUMBER: NZ 333568
 19 <151> PRIOR FILING DATE: 1998-12-23
 21 <160> NUMBER OF SEQ ID NOS: 10
 23 <170> SOFTWARE: PatentIn version 3.0
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 10
 27 <212> TYPE: PRT
 28 <213> ORGANISM: Perna canaliculus
 30 <400> SEQUENCE: 1
 32 Asp Gly Glu Gln Cys Asn Asp Gly Gln Asn
 33 1 5 10
 35 <210> SEQ ID NO: 2
 36 <211> LENGTH: 19
 37 <212> TYPE: PRT
 38 <213> ORGANISM: Perna canaliculus
 40 <400> SEQUENCE: 2
 42 Gln Gly Gly His Glu Val Glu Ser Glu Arg Val Ala Cys Cys Val Ile
 43 1 5 10 15
 45 Gly Arg Ala
 48 <210> SEQ ID NO: 3
 49 <211> LENGTH: 9
 50 <212> TYPE: PRT
 51 <213> ORGANISM: Perna canaliculus
 53 <400> SEQUENCE: 3
 55 Gly Gln Ser His Pro Glu Ile Val His
 56 1 5
 58 <210> SEQ ID NO: 4
 59 <211> LENGTH: 7
 60 <212> TYPE: PRT
 61 <213> ORGANISM: Perna canaliculus
 63 <400> SEQUENCE: 4
 65 Tyr His Gly His Asp Asp Ala
 66 1 5
 68 <210> SEQ ID NO: 5
 69 <211> LENGTH: 7
 70 <212> TYPE: PRT
 71 <213> ORGANISM: Perna canaliculus
 73 <400> SEQUENCE: 5

ENTERED

RAW SEQUENCE LISTING

DATE: 08/30/2001

PATENT APPLICATION: US/09/868,760

TIME: 09:01:49

Input Set : A:\00019900.txt

Output Set: N:\CRF3\08302001\I868760.raw

75 Val Val Asn Glu Val His His

76 1 5

78 <210> SEQ ID NO: 6

79 <211> LENGTH: 1491

80 <212> TYPE: DNA

81 <213> ORGANISM: Perna canaliculus

83 <400> SEQUENCE: 6

```

84 gayggggagc agtgtaacga tgggcagaac aaagatgacc accatgacga ccaccacgat      60
86 gatcaccatg acgaccatga tgatgatgat gaaacaatgc actatgccca gtgtgaaatg      120
88 gaaccaaacc ctcatatggc tagcagcctt caccaccatg tccatggcag catagagttg      180
90 tcacagaagg gtcattggagc tgtttatcta gaacttcata ttgtcggatt caacacaagt      240
92 gaagaccatg acgaccacca tcatggactt catctgcaca tgcttggtga catgtcagca      300
94 ggttgtgatt ctattggcga actgtacaat gctcaccacg aaaaacatgc tgaccctggt      360
96 gacctcgggt acctggttga cgatgatagg ggcgtggtta atgaagttca tcattatgct      420
98 tggttggaca ttgatgttac agcaccaaac accgaagctc tcattggaca ctcaatgact      480
100 attttacaag ggagtacacac cgatgctgat accccagcca gtagaatcgc ctggtgtgtt      540
102 attggtcatg gaaaagctcg cccagaaaca gcagctgctc tacatcacga gctagaggaa      600
104 gataaaactg agcattatgc ccattgtgac gtaagatcta atacacacca accaaaggct      660
106 cttcatcatc atgtccacgg aacctcgat ttcaaacaag ttggttatgg tgacctgaa      720
108 gtgtcctacc atttagaggg atttaatgta agtgatgacc acaaagatca tctccatgac      780
110 gtacagatct acgccaacgg tgacctgacc agtggatgtg ataacctcgg tgctaaatat      840
112 gatcctcatg aagattacca cagtgaagtg ggtgatctag gagatattca cgatgatgac      900
114 catggcgttg tcaatgaaag ccacagatat tcttgatca atatcttcgg tgatgacagt      960
116 gtcctgggac gttctattgc catcaccaa agagaccatc ttcataaaag tgccaaaatt      1020
118 gcctgttgtg tcataggacg tggacagagc catccagaaa ttgttcacag agctaaatgt      1080
120 gttgtcagac ctaatacaga atctactggt ttacatcacc atgtctctgg ttctataaca      1140
122 ttcgaacaga cccctggagg atcaacacat atgacggctg atctcaaagg atttaacgtt      1200
124 atgagagact tgtcacatca tcgtcatggt gtgcagctcc atgaatgggg agatatgtcc      1260
126 catggctgtc atctcttagg cagaatgtac catggtcagt atgatgtca tgaccccaaa      1320
128 agacctggtg accttggtga tgttatagat gattcccatg gcacgttca ttcaactaga      1380
130 acctttgatc atcttaatgt tgaagatctt aacgcacgtt cccttgatgat tatgcagggc      1440
132 ggacatgagg tgcagagtga gaggggtgct tgctgtgtta taggacgggc a          1491

```

135 <210> SEQ ID NO: 7

136 <211> LENGTH: 497

137 <212> TYPE: PRT

138 <213> ORGANISM: Perna canaliculus

140 <400> SEQUENCE: 7

```

142 Asp Gly Glu Gln Cys Asn Asp Gly Gln Asn Lys Asp Asp His His Asp
143 1 5 10 15
145 Asp His His Asp Asp His His Asp Asp His Asp Asp Asp Asp Glu Thr
146 20 25 30
148 Met His Tyr Ala Gln Cys Glu Met Glu Pro Asn Pro His Met Ala Ser
149 35 40 45
151 Ser Leu His His His Val His Gly Ser Ile Glu Leu Ser Gln Lys Gly
152 50 55 60
154 His Gly Ala Val Tyr Leu Glu Leu His Leu Val Gly Phe Asn Thr Ser
155 65 70 75 80
157 Glu Asp His Asp Asp His His His Gly Leu His Leu His Met Leu Gly
158 85 90 95

```

RAW SEQUENCE LISTING

DATE: 08/30/2001

PATENT APPLICATION: US/09/868,760

TIME: 09:01:49

Input Set : A:\00019900.txt

Output Set: N:\CRF3\08302001\I868760.raw

```

160 Asp Met Ser Ala Gly Cys Asp Ser Ile Gly Glu Leu Tyr Asn Ala His
161      100      105      110
163 Pro Glu Lys His Ala Asp Pro Gly Asp Leu Gly Asp Leu Val Asp Asp
164      115      120      125
166 Asp Arg Gly Val Val Asn Glu Val His His Tyr Ala Trp Leu Asp Ile
167      130      135      140
169 Asp Gly Thr Ala Pro Asn Thr Glu Ala Leu Ile Gly His Ser Met Thr
170 145      150      155      160
172 Ile Leu Gln Gly Ser His Thr Asp Ala Asp Thr Pro Ala Ser Arg Ile
173      165      170      175
175 Ala Cys Cys Val Ile Gly His Gly Lys Ala Arg Pro Glu Thr Ala Ala
176      180      185      190
178 Ala Leu His His Glu Leu Glu Glu Asp Lys Thr Glu His Tyr Ala His
179      195      200      205
181 Cys Asp Val Arg Ser Asn Thr His Gln Pro Lys Ala Leu His His His
182      210      215      220
184 Val His Gly Thr Ile Asp Phe Lys Gln Val Gly Tyr Gly Asp Leu Glu
185 225      230      235      240
187 Val Ser Tyr His Leu Glu Gly Phe Asn Val Ser Asp Asp His Lys Asp
188      245      250      255
190 His Leu His Asp Val Gln Ile Tyr Ala Asn Gly Asp Leu Thr Ser Gly
191      260      265      270
193 Cys Asp Asn Leu Gly Ala Lys Tyr Asp Pro His Glu Asp Tyr His Ser
194      275      280      285
196 Glu Leu Gly Asp Leu Gly Asp Ile His Asp Asp Asp His Gly Val Val
197      290      295      300
199 Asn Glu Ser His Arg Tyr Ser Trp Ile Asn Ile Phe Gly Asp Asp Ser
200 305      310      315      320
202 Val Leu Gly Arg Ser Ile Ala Ile His Gln Arg Asp His Leu His Lys
203      325      330      335
205 Ser Ala Lys Ile Ala Cys Cys Val Ile Gly Arg Gly Gln Ser His Pro
206      340      345      350
208 Glu Ile Val His Arg Ala Lys Cys Val Val Arg Pro Asn Thr Glu Ser
209      355      360      365
211 Thr Gly Leu His His His Val Ser Gly Ser Ile Thr Phe Glu Gln Thr
212      370      375      380
214 Pro Gly Gly Ser Thr His Met Thr Ala Asp Leu Lys Gly Phe Asn Val
215 385      390      395      400
217 Ser Glu Asp Leu Ser His His Arg His Gly Val Gln Leu His Glu Trp
218      405      410      415
220 Gly Asp Met Ser His Gly Cys His Ser Leu Gly Arg Met Tyr His Gly
221      420      425      430
223 His Asp Asp Ala His Asp Pro Lys Arg Pro Gly Asp Leu Gly Asp Val
224      435      440      445
226 Ile Asp Asp Ser His Gly Ile Val His Ser Thr Arg Thr Phe Asp His
227      450      455      460
229 Leu Asn Val Glu Asp Leu Asn Ala Arg Ser Leu Val Ile Met Gln Gly
230 465      470      475      480
232 Gly His Glu Val Glu Ser Glu Arg Val Ala Cys Cys Val Ile Gly Arg

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,760

DATE: 08/30/2001

TIME: 09:01:49

Input Set : A:\00019900.txt

Output Set: N:\CRF3\08302001\I868760.raw

```

233                               485                               490                               495
235 Ala
238 <210> SEQ ID NO: 8
239 <211> LENGTH: 1611
240 <212> TYPE: DNA
241 <213> ORGANISM: Perna canaliculus
243 <220> FEATURE:
244 <221> NAME/KEY: misc_feature
245 <222> LOCATION: (1)..(1611) 0 ✓
246 <223> OTHER INFORMATION: 'n' can be any one of the nucleotides 'a', 'c', 'g' or 't';
249 <400> SEQUENCE: 8
250 gaygggggagc agtgtaacga tgggcagaac aaagatgacc accatgacga ccaccacgat      60
252 gatcaccatg acgaccatga tgatgatgat gaaacaatgc actatgccca gtgtgaaatg      120
254 gaaccaaacc ctcatatggc tagcagcctt caccaccatg tccatggcag catagagttg      180
256 tcacagaagg gtcattggagc tgtttatcta gaacttcac ttgtcggatt caacacaagt      240
258 gaagaccatg acgaccacca tcatggactt catctgcaca tgcttggtga catgtcagca      300
260 ggtttgtgatt ctattggcga actgtacaat gctcaccag aaaaacatgc tgaccctggt      360
262 gacctcggtg acctggttga cgatgatagg ggcgtggtta atgaagttca tcattatgct      420
264 tggtttgaca ttgatgttac agcaccaaacc accgaagctc tcattggaca ctcaatgact      480
266 attttacaag ggagtacac cgatgtgatg accccagcca gtagaatcgc ctggttggtt      540
268 attggtcatg gaaaagctcg cccagaaaca gcagctgctc tacatcacga gctagaggaa      600
270 gataaaactg agcattatgc ccattgtgac gtaagatcta atacacacca accaaaggct      660
272 cttcatcatc atgtccacgg aacctcgat ttcaacaag ttggttatgg tgaccttgaa      720
274 gtgtcctacc atttagaggg atttaatgta agtcatgacc acaaagatca tctccatgac      780
276 gtacagatct acgccaacgg tgacctgacc agtggatgtg ataacctcg tgctaaatat      840
278 gatcctcatg aagattacca cagtgaagtg ggtgatctag gagatattca cgatgatgac      900
280 catggcgttg tcaatgaaag ccacagatat tcttgatca atatcttcgg tgatgacagt      960
282 gtccctgggac gttctattgc cattcaccaa agagaccatc ttcataaaag tgccaaaatt      1020
284 gcctgttgtg tcataggacg tggacagagc catccagaaa ttgttcacag agctaaaatg      1080
286 gttgtcagac ctaatacaga atctactggg ttacatcacc atgtctctgg ttctataaca      1140
288 ttccaacaga ccctggagg atcaacacat atgacggctg atctcaaagg atttaacgtt      1200
290 agtgaggact tgtcacatca tcgtcatggt gtgcagctcc atgaatgggg agatatgtcc      1260
292 catggctgtc actccttagg cagaatgtac catggtcatg atgatgctca tgaccccaaa      1320
294 agacctgggt accttggtga tgttatagat gattcccatg gcacgttca ttcaactaga      1380
296 acctttgatc atcttaatgt tgaagatctt aacgcacggt cccttggtgat tatgcagggc      1440
298 ggacatgagg tcgagagtga gaggggtgct tgctgtgtta taggacgggc atgaataacc      1500
300 tcaactagagt gactttgtct aacatgacaa ttaacaattg tataacttcg ctaaaaaata      1560
W--> 302 aaacaatgac acaatgnaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a      1611
305 <210> SEQ ID NO: 9
306 <211> LENGTH: 8
307 <212> TYPE: PRT
308 <213> ORGANISM: terrestrial leech
310 <400> SEQUENCE: 9
312 Gly Gln Ser Cys Asn Asp Gly Gln
313 1                               5
315 <210> SEQ ID NO: 10
316 <211> LENGTH: 29
317 <212> TYPE: DNA
318 <213> ORGANISM: Artificial Sequence

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/868,760

DATE: 08/30/2001

TIME: 09:01:50

Input Set : A:\00019900.txt

Output Set: N:\CRF3\08302001\I868760.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

L:331 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10



Creation date: 11-20-2003
Indexing Officer: ICHARLES - IRENE CHARLES
Team: OIPEBackFileIndexing
Dossier: 09868760

Legal Date: 11-26-2001

No.	Docode	Number of pages
1	M903	2

Total number of pages: 2

Remarks:

Order of re-scan issued on